

Opportunistic Routing with Congestion Diversity in Wireless Ad-hoc Networks

A. A. Bhorkar, M. Naghshvar, T. Javidi

Department of Electrical Engineering,

University of California San Diego, La Jolla, CA 92093

abhorkar@ucsd.edu, naghshvar@ucsd.edu, tjavidi@ucsd.edu

Abstract—We consider the problem of routing packets across a multi-hop network consisting of multiple sources of traffic and wireless links while ensuring bounded expected delay. Each packet transmission can be overheard by a random subset of receiver nodes among which the next relay is selected opportunistically. The main challenge in the design of minimum-delay routing policies is balancing the trade-off between routing the packets along the shortest paths to the destination and distributing traffic according to the maximum backpressure. Combining important aspects of shortest path and backpressure routing, this paper provides a systematic development of a distributed opportunistic routing policy with congestion diversity (D-ORCD).

D-ORCD uses a measure of draining time to opportunistically identify and route packets along the paths with an expected low overall congestion. D-ORCD is proved to ensure a bounded expected delay for all networks and under any admissible traffic. Furthermore, this paper proposes a practical implementation which empirically optimizes critical algorithm parameters and their effects on delay as well as protocol overhead. Realistic Qualnet simulations for 802.11-based networks demonstrate a significant improvement in the average delay over comparative solutions in the literature.

Index Terms—wireless, ad-hoc networks, routing, congestion, implementation

I. INTRODUCTION

Opportunistic routing for multi-hop wireless ad-hoc networks has long been proposed to overcome deficiencies of conventional routing [1]–[5]. Opportunistic routing mitigates the impact of poor wireless links by exploiting the broadcast nature of wireless transmissions and the path diversity. More precisely, the routing decisions are made in an online manner by choosing the next relay based on the actual transmission outcomes as well as a rank ordering of neighboring nodes.

The authors in [4] provided a Markov decision theoretic formulation for opportunistic routing and a unified framework for many versions of opportunistic routing [1]–[3], with the variations due to the authors' choices of costs. In particular, it is shown that for any packet, the optimal routing decision, in the sense of minimum cost or hop-count, is to select the next relay node based on an index. This index is equal to the expected cost or hop-count of relaying the packet along the least costly or the shortest feasible path to the destination. When multiple streams of packets are to traverse the network, however, it might be necessary to route some packets along longer or more costly paths, if these paths eventually lead to links that are less congested. More precisely,

and as noted in [6], [7], the opportunistic routing schemes in [1]–[5] can potentially cause severe congestion and unbounded delays (see examples given in [6]). In contrast, it is known that an opportunistic variant of backpressure [8], diversity backpressure routing (DIVBAR) [7] ensures bounded expected total backlog for all stabilizable arrival rates. To ensure throughput optimality (bounded expected total backlog for all stabilizable arrival rates), backpressure-based algorithms [7], [8] do something very different from [1]–[5]: rather than using any metric of closeness (or cost) to the destination, they choose the receiver with the largest positive differential backlog (routing responsibility is retained by the transmitter if no such receiver exists). This very property of ignoring the cost to the destination, however, becomes the bane of this approach, leading to poor delay performance in low to moderate traffic (see [6]). Other existing provably throughput optimal routing policies [9]–[12] distribute the traffic locally in a manner similar to DIVBAR and, hence, result in large delay.

Recognizing the shortcomings of the two approaches, researchers have begun to propose solutions which combine elements of shortest path and backpressure computations [7], [15], [16]. In [7], E-DIVBAR is proposed: when choosing the next relay among the set of potential forwarders, E-DIVBAR considers the sum of the differential backlog and the expected hop-count to the destination (also known as ETX). However, as shown in [6], E-DIVBAR does not necessarily result in a better delay performance than DIVBAR. Instead of a simple addition used in EDIVBAR, this paper provides a distributed opportunistic routing policy with congestion diversity (D-ORCD) under which the congestion information is integrated with the distributed shortest path computations of [4]. In our previous work [13], ORCD, a centralized version of D-ORCD, is shown to be throughput optimal without discussion on system implications. In this paper, we extend the throughput optimality proof for the distributed version and discuss implementation issues in detail. We also tackle some of the system level issues observed in realistic settings via detailed Qualnet simulations. We then show that D-ORCD exhibits better delay performance than state of the art routing policies, namely, EXOR, DIVBAR and E-DIVBAR.

Before we close, we emphasize that some of the ideas behind the design of D-ORCD have also been used as guiding principles in many routing solutions: some in opportunistic context [14], [15] and some in conventional context [16].

Below, we detail the similarity and differences between these solutions and our work for the sake of completeness, even though, in our study, we have chosen to focus only on solutions with comparable overhead and similar degree of practicality. In [14], perhaps the most related work to ours, the authors consider a flow-level model of the network and propose a routing policy referred to as *min-backlogged-path* routing, under which the flows are routed along the paths with minimum total backlog. In this light, D-ORCD can be viewed as a packet-based version of the min-backlogged-path routing without a need for the enumeration of paths across the network and costly computations of total backlog along paths. In [15], authors propose a modified version of backpressure which uses the shortest path information to minimize the average number of hops per packet delivery, while keeping the queues stable. In [16], a modified throughput optimal backpressure policy, LIFO-Backpressure, is proposed using LIFO discipline at layer 2. Neither of these approaches lend themselves to practical implementations: [15] requires maintaining large number of virtual queues at each node increasing implementation complexity, while [16] uses atypical LIFO scheduler resulting in significant reordering of packets. Furthermore, while LIFO-Backpressure policy guarantees stability with minimal queue-length variations, realistic bursty traffic in large multi-hop wireless networks may result in queue-length variations and unnecessarily high delay.

The paper is organized as follows. In Section II, we describe the D-ORCD routing algorithm. In Section III, we discuss various protocol implementation issues of D-ORCD. Section IV describes our simulation results in detail, where we compare the performance of various routing policies with D-ORCD. We then discuss theoretical guarantees of D-ORCD in Section V. We provide concluding remarks and discuss directions for future research work in Section VI. The appendix contains proofs of throughput optimality of D-ORCD under certain assumptions on the model.

II. OPPORTUNISTIC ROUTING WITH CONGESTION DIVERSITY

The goal of this paper is to design a routing policy with improved delay performance over existing opportunistic routing policies. In this section, we describe the guiding principle and the design of Opportunistic Routing with Congestion Diversity (D-ORCD). We propose a time-varying distance vector, which enables the network to route packets through a neighbor with the least estimated delivery time.

D-ORCD opportunistically routes a packet using three stages of: (a) transmission, (b) acknowledgment, and (c) relaying. During the transmission stage, a node transmits a packet. During the acknowledgment stage, each node that has successfully received the transmitted packet, sends an acknowledgment (ACK) to the transmitter node. D-ORCD then takes routing decisions based on a congestion-aware distance vector metric, referred to as the *congestion measure*. More specifically, during the relaying stage, the relaying responsibility of the packet is shifted to a node with the least congestion measure among the ones that have received the

TABLE I
NOTATIONS USED IN THE DESCRIPTION OF THE ALGORITHM

Symbol	Definition
$\mathcal{N}(i)$	Neighbours of node i
$V_i^d(t)$	Congestion measure at node i at time t
$\tilde{V}_k^{(i,d)}(t)$	Congestion measure obtained at node i from k
$T(t)$	Ending time of the latest computation cycle before time t
T_c	Duration of the computation interval
T_s	Control packet transmission interval
$L_i(t)$	Local congestion at node i
$D_i(t)$	Congestion down the stream for node i
$K_{D-ORCD}^{(i,d)}(t)$	Selected relay for transmission at node i
$S_i(t)$	Set of nodes receiving packet transmitted by node i
$Q_i^d(t)$	Queue-length at node i destined for d at time t
$\bar{Q}_i^d(t)$	Average queue-length at node i destined for d
$P_{succ-k}^{(i,d)}(t)$	Probability that highest priority node k receives packet
$P^{(i,d)}(t)$	Probability that at-least one higher priority node receives packet
$H^{(i,d)}(t)$	Set of higher priority nodes than node i

packet. The congestion measure of a node associated with a given destination provides an estimate of the best possible draining time of a packet arriving at that node until it reaches destination. Each node is responsible to update its congestion measure and transmit this information to its neighbors. Next, we detail D-ORCD design and the computations performed at each node to update the congestion measure.

A. D-ORCD Design

We consider a network of D nodes labelled by $\Omega = \{1, \dots, D\}$. Let p_{ij} be the probability that the packet transmitted by node i is successfully received by node j . Node j is said to be *reachable* by node i , if $p_{ij} > 0$. The set of all nodes in the network which are reachable by node i is referred to as *neighborhood* of node i and is denoted by $\mathcal{N}(i)$.

D-ORCD relies on a routing table at each node to determine the next best hop. The routing table at node i consists of a list of neighbors $\mathcal{N}(i)$ and a structure consisting of estimated congestion measure for all neighbors in $\mathcal{N}(i)$ associated with different destinations. The routing table acts as a storage and decision component at the routing layer. The routing table is updated using a “virtual routing table” at the end of every “computational cycle”: an interval T_c units of time. To update virtual routing table, during the progression of the computation cycle the nodes exchange and compute the temporary congestion measures. The temporary congestion measures are computed in a fashion similar to a distributed stochastic routing computation of [4] using the backlog information at the beginning of the computation cycle (generalizing the

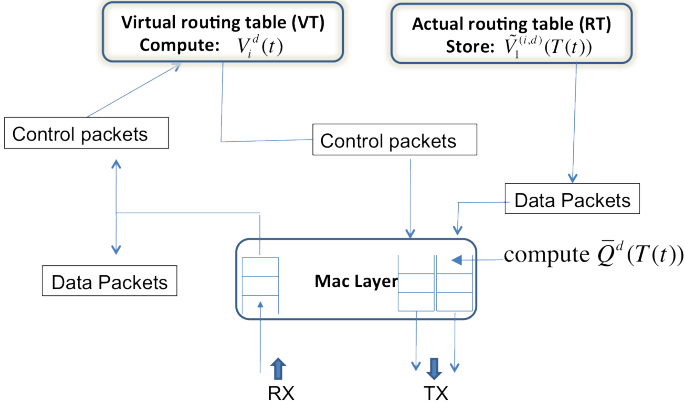


Fig. 1. Operation of D-ORCD

computations of distributed Bellman-Ford). We conceptualize this in terms of a virtual routing table updating and maintaining these temporary congestion measures. We assume that each node has a common global time to ensure that the nodes update the routing table roughly at the same time.

We denote the temporary congestion measure associated with node $i \in \Omega$ at time t and destination $d \in \Omega$ as $V_i^d(t)$. Each node i computes $V_i^d(t)$ based on congestion measures $\tilde{V}_k^{(i,d)}(t)$ obtained via periodic communication with its neighbours $k \in \mathcal{N}(i)$ and the queue backlog at the start of the computation cycle. D-ORCD stores these temporary congestion measures $\{V_i^d(t)\}_{d \in \Omega}$ and $\{\tilde{V}_k^{(i,d)}(t)\}_{d \in \Omega, k \in \mathcal{N}(i)}$ in the virtual routing table. More precisely, node i periodically compute its own congestion measure and subsequently advertises it to its neighbors using control packets at intervals of $T_s \leq T_c$ seconds. Finally the actual routing table is updated using the entries in the virtual routing table after every T_c seconds. The sequence of operations performed by D-ORCD are shown in Figs. 1,2.

Meanwhile, for routing decisions, node i uses the entries in the actual routing table (computed during the last computation cycle). Let $T(t) = \max_n \{nT_c : nT_c \leq t, n \in \mathbb{Z}\}$ be the ending time of the latest computation cycle. Then node i stores $\tilde{V}_k^{(i,d)}(T(t))$ in the actual routing table and selects the next best hop $K_{D-ORCD}^{(i,d)}$ to minimize the packet's draining time, i.e.

$$K_{D-ORCD}^{(i,d)}(t) = \arg \min_{k \in S_i(t) \cup i} \tilde{V}_k^{(i,d)}(T(t)), \quad (1)$$

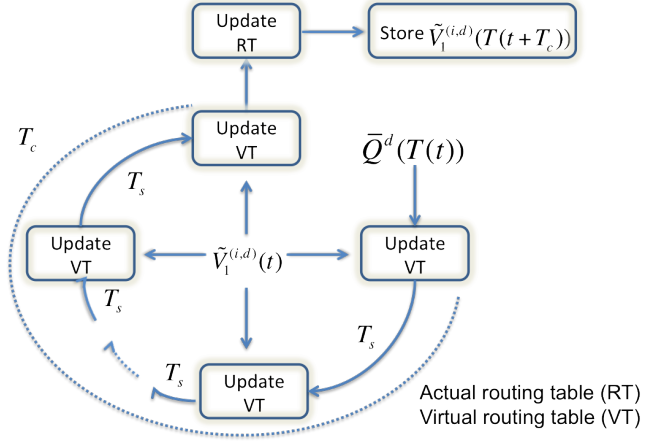
where $S_i(t)$ denotes a random set of nodes receiving the packet transmitted by node i at time t .

Next, we describe the distributed computations performed during each computation cycle.

B. Congestion Measure Computations

The congestion measure associated with node i for a destination d at time t is the aggregate sum of the local draining time at node i (denoted by $L_i^d(t)$) and the draining time from its next hop to the destination (denoted by $D_i^d(t)$), i.e.

$$V_i^d(t) = L_i^d(t) + D_i^d(t). \quad (2)$$

Fig. 2. Actual routing table is updated every T_c units of time while virtual routing table is updated after receiving any control packet

Assuming a FIFO discipline at layer-2, we proceed to decompose the local draining time. This relies on the observation that when a packet arrives at a node, i , its waiting time is equal to the time spent in draining the packets that have arrived earlier plus its own transmission time. If $P^{(i,d)}(t)$ denotes the probability that the packet transmitted by node i is successfully received by a node with lower congestion measure, then expected transmission time at node i for the packet is given by $\frac{1}{P^{(i,d)}(t)}$. Let $\bar{Q}_i^d(t)$ denote the number of packets destined for destination d averaged over previous computation cycle. $\bar{Q}_i^d(t)$ is updated as

$$\bar{Q}_i^d(t) = \frac{T_s}{T_c} \sum_{l=0}^{\frac{T_c}{T_s}-1} Q_i^d(T(t) - l).$$

The local draining time for node i to destination d at time t is approximated as,

$$L_i^d(t) = \frac{1}{P^{(i,d)}(t)} + \sum_{d' \in \Omega} \frac{\bar{Q}_i^{d'}(T(t))}{P^{(i,d')}(t)}. \quad (3)$$

D-ORCD computes the expected congestion measure “down the stream” for each node $i \in \Omega$ using the latest congestion measure $\tilde{V}_k^{(i,d)}(t)$ received from nodes $k \in \Omega$ with lower congestion measure. With respect to the destination d , a node $k \in \Omega$ is defined as a higher priority node than node i if $\tilde{V}_k^{(i,d)}(t) < V_i^d(t)$ and the set of higher priority nodes as $H^{(i,d)}(t)$. Let $P_{succ-k}^{(i,d)}(t)$ be the probability that node k is the highest priority node to successfully hear node i at time t and $k \in H^{(i,d)}(t)$. As a result, the expected congestion “down the stream” $D_i^d(t)$ can be given as

$$D_i^d(t) = \sum_{k \in \Omega} P_{succ-k}^{(i,d)}(t) \tilde{V}_k^{(i,d)}(t). \quad (4)$$

Remark 1. In each computation cycle, assuming T_c is large, D-ORCD computations converge to the Bellman equation associated with the minimum cost (“shortest path”) route in a network, where the link costs are given in terms of the queue length $\bar{Q}_i^d(t)$.

Remark 2. If the links success probabilities have independent realizations, then for all $S \subseteq \Omega$, $P(S_i(t) = S) = \prod_{k \in S} \prod_{l \notin S} P_{ik}(1 - p_{il})$. The success probabilities $P^{(i,d)}(t)$ and $P_{succ-k}^{(i,d)}(t)$ can be calculated as

$$P^{(i,d)}(t) = \sum_{S: \exists k \in S \cap H^{(i,d)}(t)} P(S_i(t) = S), \quad (5)$$

$$P_{succ-k}^{(i,d)}(t) = \frac{1}{P^{(i,d)}(t)} \times \sum_{\substack{S: \tilde{V}^{(k,d)}(t) < \tilde{V}^{(k',d)}(t) \\ k', k \in S \cap H^{(i,d)}(t)}} P(S_i(t) = S). \quad (6)$$

C. Opportunistic Routing with Partial Diversity

The three-way handshake procedure discussed in Section II-A to achieve receiver diversity gain in an opportunistic scheme is achieved at the cost of an increase in the control overhead. In particular, it is easy to see that this overhead cost, which is the total number of ACKs sent per data packet transmission, increases linearly with the size of the set of potential forwarders. Thus, we consider a modification of D-ORCD in the form of opportunistically routing with partial diversity (P-ORCD). This class of routing policies is parametrized by parameter M denoting the maximum number of forwarder nodes. This is equivalent to a constraint on the maximum number of nodes allowed to send acknowledgment per data packet transmission. Such a constraint will sacrifice the diversity gain, and hence the performance of any opportunistic routing policy, for lower communication overhead.

In order to implement opportunistic routing policies with partial diversity, before the transmission stage occurs we find the set of “best neighbors” for each node i at any time t , denoted by $B_i^*(t)$, where $|B_i^*(t)| \leq M$. After transmission of a packet from node i at time t , the routing decision is made as follows: 1) among the nodes in $B_i^*(t) \cap S_i(t)$, select a node with the lowest congestion measure as the next forwarder; or 2) retain the packet if none of the nodes in the set $B_i^*(t)$ has received the packet. Next we give a mathematical formulation for modifications of D-ORCD with partial diversity.

Let \mathcal{B} be the collection of all subsets of Ω of size less than or equal to M , i.e. $\mathcal{B} = \{B \subseteq \Omega : |B| \leq M\}$.

In D-ORCD protocol with partial diversity, (PD-ORCD), the corresponding quantities $\tilde{V}_i^d(t)$ are updated as

$$\tilde{V}_i^d(t) = \min_{B \in \mathcal{B}} \left\{ L_i^d(t) + \sum_{k: k \in B} P_{succ-k}^{(i,d)}(t) \tilde{V}_k^{(i,d)}(t) \right\}, \quad (7)$$

while the next hop is selected as

$$K_{PD-ORCD}^{(i,d)}(t) = \arg \min_{k \in \{S_i(t) \cap B\} \cup i} \tilde{V}_k^{(i,d)}(T(t)). \quad (8)$$

We carry out a simulation study for the delay performance of D-ORCD with these modifications and compare it to the delay performance of the other routing policies in Section IV.

Remark 3. When $M = 1$ each node can send packets only to one of its neighbors. Therefore, this routing policy cannot take the advantage of the broadcast nature of wireless transmissions anymore, and is classified as conventional routing.

In the next section, we discuss the practical issues associated with computation of the time-varying congestion measures $V_i^d(t)$, $i \in \Omega$. Furthermore, we propose practical implementations and heuristics.

III. IMPLEMENTATION DETAILS: PROTOCOL COMPONENTS

In this section we discuss the implementation issues of D-ORCD which involves distributed and asynchronous iterative computations of $V_i^d(t)$'s. We provide brief discussion of the basic challenges of D-ORCD including the three-way handshake procedure employed at MAC, link quality estimation, and avoidance of loops while routing.

A. 802.11 Compatible Implementation

1) *Three way Handshake:* The implementation of D-ORCD, analogous to any opportunistic routing scheme, involves the selection of a relay node among the candidate set of nodes that have received and acknowledged a packet successfully. One of the major challenges in the implementation of an opportunistic routing algorithm, in general, and D-ORCD in particular, is the design of an 802.11 compatible acknowledgement mechanism at the MAC layer. Below we propose a practical and simple way to implement acknowledgement architecture.

The transmission at any node i is done according to 802.11 CSMA/CA mechanism. Specially, before any transmission, transmitter i performs channel sensing and starts transmission after the backoff counter is decremented to zero. For each neighbor node $j \in \mathcal{N}(i)$, the transmitter node i then reserves a virtual time slot of duration $T_{ACK} + T_{SIFS}$, where T_{ACK} is the duration of the acknowledgement packet and T_{SIFS} is the duration of Short Inter Frame Space (SIFS) [17]. Transmitter i then piggy-backs a priority ordering of nodes $\mathcal{N}(i)$ with each data packet transmitted. The priority ordering determines the virtual time slot in which the candidate nodes transmit their acknowledgement. Nodes in the set S_i that have successfully received the packet then transmit acknowledgement packets sequentially in the order determined by the transmitter node.

After a waiting time of $T_{wait} = |\mathcal{N}(i)|(T_{ACK} + T_{SIFS})$ during which each node in the set S_i has had a chance to send an ACK, node i transmits a FOrwarding control packet (FO). The FO packets contain the identity of the next forwarder, which may be node i itself (i.e. node i retains the packet) or any node $j \in S_i$. If T_{wait} expires and no FO packet is received (FO packet reception is unsuccessful), then the corresponding candidate nodes drop the received data packet. If transmitter i does not receive any acknowledgement, it retransmits the packet. The backoff window is doubled after every retransmission. Furthermore, the packet is dropped if the retry limit (set to 7) is reached.

B. Control Packets fidelity

D-ORCD depends on a reliable, frequent, and timely delivery of the control packets. As documented in [18], the loss of control packets may destabilize the algorithm operation and cause significant performance degradation for many well

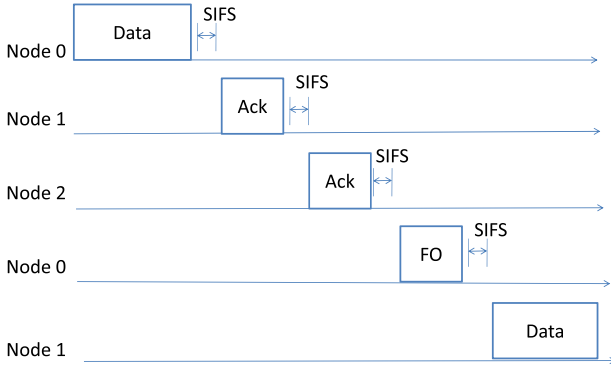


Fig. 3. Typical packet transmission sequence for D-ORCD when packet transmitted by node 0 is received by nodes 1 and 2.

known routing algorithms. In our implementation, we have taken advantage of the priority-based queuing to implement this component of the control plane. D-ORCD prioritizes the control packets by assigning them the highest strict priority, reducing the probability that the packets are dropped at the MAC layer and also ensuring a timely delivery of the control packets. In particular, D-ORCD utilizes priority queues: data packets are assigned to the lower priority queue and control packets are assigned to the higher priority queue. Moreover, D-ORCD scheduler assigns a sufficiently lower PHY rate for the control packets.

C. Link Quality Estimation Protocol

D-ORCD computations given by (2) utilize link success probabilities p_{ij} for each pair of nodes i, j . We now describe method to determine the probability of successfully receiving a data packet for each pair of nodes $i, j \in \Omega$. It consists of two phases: active probing and passive probing. In the active probing, dedicated probe packets are broadcasted periodically to estimate link success probabilities. In passive probing, we utilize the overhearing capability of the wireless medium. The nodes are configured to promiscuous mode, hence enabling them to hear the packets from neighbors. In passive probing, the MAC layer keeps track of the number of packets received from the neighbors including the retransmissions. Finally, a weighted average is used to combine the active and passive estimates to determine the link success probabilities. Passive probing does not introduce any additional overhead cost but can be slow, while active probing rate is set independently of the data rate but introduces costly overhead.

D. Loop Avoidance Heuristic

D-ORCD approximates the solution to the fixed point equation via distributed distance vector approach. The classical problem of counting to infinity [19] in distance vector routing can affect D-ORCD performance due to the time varying nature of the congestion metric. The problem is most acute

when there is a sudden burst of traffic.¹ and can cause severe transient effects due to slow updates of the control packets. The looping results in large delays, increased interference and loss of packets.²

To address this issue, in our experiments we utilize an extension of the Split-horizon with poison reverse solution [20] to avoid loops. In Split-horizon with poison reverse, a node advertises routes as unreachable to the node through which they were learned. We have extended the rule to D-ORCD by advertising the routes as unreachable to higher ranked nodes. This removes most looping routes before they can propagate through the network.

IV. SIMULATIONS

In this section, we compare the expected delay encountered by the packets in the network under various opportunistic routing policies: ExOR, DIVBAR, E-DIVBAR and D-ORCD in Qualnet simulations. We first investigate the performance of D-ORCD with respect to a canonical example to demonstrate D-ORCD gains [6]. We then use a realistic topology of 16 nodes placed in a grid topology to demonstrate the robust performance improvement in practical settings.

A. The Simulation Setup

Our simulations are performed in QualNet. We consider two set of topologies in our experimental study:

- 1) Canonical Example: In this example, we study the canonical example in Fig. 4. We motivate the performance improvement for D-ORCD by a scenario which exemplify the need to avoid congestion in the network by highlighting the shortcomings of the existing routing paradigms: shortest path and backpressure.
- 2) Grid Topology: We study an outdoor wireless settings of grid topology consisting of 16 nodes separated by a distance of 200 meters. These simulations demonstrate a robust performance gain under D-ORCD in a realistic network.

We now describe the parameters settings in the simulation. The nodes are equipped with 802.11b radios transmitting at 11 Mbps with transmission power 15 dBm. The wireless medium model includes Rician fading with K-factor of 4 and Log-normal shadowing with mean 4dB. In the canonical example path loss is determined by pathloss matrix which gives the attenuation of the received signal power with distance from the transmitter for every pair of network nodes, while for grid topology the path loss follows ITM model in [21]. The antenna model is the standard omnidirectional antenna model with the default settings of the simulator. The network queues are FIFO with finite buffer of 750 KB.

The acknowledgement packets are short packets of length 24 bytes transmitted at 11 Mbps, while FO packets are of length 20 bytes and transmitted at lower rate of 1 Mbps to ensure reliability. If unspecified, packets are generated

¹Similar to the broken link scenario in a typical distance vector routing.

²Packet loss occurs when time to live (TTL) value exceeds the number of allowed hops (typically 64)

according to a poisson modulated Markov traffic. The packets are assumed to be of length 512 bytes equipped with simple cyclic redundancy check (CRC) error detection. The control packets are transmitted periodically at an interval of $T_s = 0.5$ seconds.

We have chosen partial diversity $M = 4$ and update frequency $T_c = T_s = 0.5$ seconds in our experimentations. A discussion on the choice of parameters in the design of D-ORCD is provided in Section IV-D.

In our study, we have compared the performance of D-ORCD against the state of the art routing algorithms. Before we proceed, we describe these candidate algorithms as well as our implementation of them.

- **DIVBAR [22]:** We implemented DIVBAR to select the next hop based on a weighted differential backlog. Specifically, let $\tilde{Q}_k^{(i,d)}(t)$ denote the latest information at node i about the number of packets buffered in queue k destined for destination d . For any destination d , DIVBAR chooses the next hop $K_{DIVBAR}^{(i,d)}(t)$, such that

$$K_{DIVBAR}^{(i,d)}(t) = \arg \min_{k \in S_i(t) \cup i} (\tilde{Q}_k^{(i,d)}(t) - Q_i^d(t)). \quad (9)$$

We have created virtual queues for each destination to identify differential backlog associated with different destinations. Note that original backpressure algorithm proposed in [22] is done in conjunction with a scheduler to maximize the network's overall weighted differential backlog as well as a mechanism to choose destination queue to be served. In our implementation, we serve the packets in a prioritized manner based on the destination using 802.11 MAC. Specifically, packet with destination $m(t)$ is selected among all possible virtual queues such that

$$m(t) = \arg \min_d \{ \min_k (\tilde{Q}_k^{(i,d)}(t) - Q_i^d(t)) \}. \quad (10)$$

In order to implement a priority scheduling we utilize a priority scheduler such that the packet destined for $m(t)$ is assigned higher priority queue. We have implemented the DIVBAR algorithm using a structure similar to D-ORCD (in which $V_i^d(t)$ is replaced with $Q_i^d(t)$).

- **ExOR [3] :** ExOR uses ETX metric when routing the packet without considering queuing information at the nodes. Specifically, for a packet destined for node d , the next hop $K_{ExOR}^{(i,d)}$ is chosen such that

$$K_{ExOR}^{(i,d)}(t) = \arg \min_{k \in S_i(t) \cup i} ETX^{(k,d)}, \quad (11)$$

where $ETX^{(k,d)}$ is the minimum number of transmissions from node k to destination d given by,

$$ETX^{(k,d)} = \min_j \left\{ \frac{1}{p_{kj}} + ETX^{(j,d)} \right\}. \quad (12)$$

We have used our distributed architecture for the calculation of ETX metric by taking $Q_i(t) = 1$ for all $i \in \Omega$ and $M = 1$ in the calculation of $V_i(t)$, even though, in principle, the overhead can be held much lower due to the time invariant nature of node ordering.

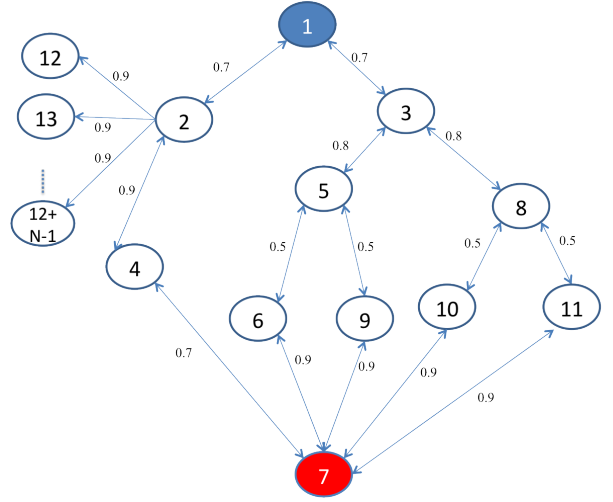


Fig. 4. Structure of the canonical network from [6]. The fractions on the links show the probability of successful transmission on each link.

- **E-DIVBAR [7]:** E-DIVBAR is a variant of DIVBAR, where along with the queue information, ETX metric is used for path selection. In particular, for a packet destined for d , the next hop $K_{E-DIVBAR}^{(i,d)}$ is chosen such that

$$K_{E-DIVBAR}^{(i,d)}(t) = \arg \min_{k \in S_i(t) \cup i} \left\{ (\tilde{Q}_k^{(i,d)}(t) - Q_i^d(t)) + ETX^{(k,d)} \right\}. \quad (13)$$

E-DIVBAR algorithm is also implemented using a structure identical to D-ORCD and DIVBAR, however, the control packets contain information about the queue-length as well as the ETX for a given destination. The commodity selection is performed using the same equation (10) as DIVBAR.

Next, we study the canonical example where we compare the average delay encountered by packets in the network under various routing policies: ExOR, DIVBAR, E-DIVBAR and D-ORCD. The choice of the canonical network enables us to clearly reveal the high capability of D-ORCD in balancing the traffic taking advantage of path diversity in the network.

B. Performance of D-ORCD: Canonical Example

Consider the network shown in Fig. 4 which is parameterized by N . Nodes 12, 13, \dots , $12 + (N - 1)$ form a “hole” in the network whose size is controlled by the parameter N . We now discuss the delay gains under D-ORCD as parameters N and λ_1 (the incoming traffic rate at node 1) are varied and verify them in this section.

Note that the source node 1 can route packets either through node 2 or node 3. Since only node 1 has a routing choice we focus on the delay experienced by packets originating in node 1. Fig. 5 provide plots of the average end-to-end packet delay and the buffer overflow ratios for all the routing algorithms as the arrival rate λ_1 is varied. We observe that D-ORCD has better delay performance than the other algorithms over the range of incoming traffic rates considered. Fig. 6 plots

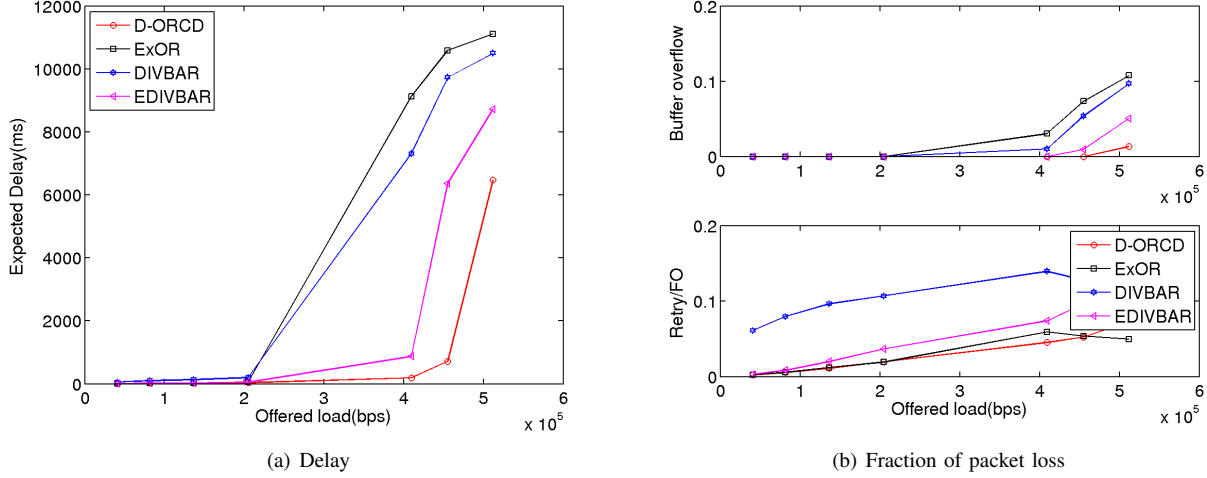


Fig. 5. Performance for Canonical Example for $N=2$

the highest priority next hop for node 1 under the candidate protocols throughout the duration of the experiments.

ExOR gives higher priority to node 2 than node 3 independent of the congestion at intermediate nodes ($ETX^{(2,7)} = 2.53$ and $ETX^{(3,7)} = 4.36$). ExOR can thus suffer from poor delay performance as the arrival rate at node 2 approaches capacity. ExOR has the worst delay performance among all the algorithms as seen in Fig 6 particularly when the traffic load on the network is high. In Fig. 6 we observe that DIVBAR and E-DIVBAR forward significant number of packets into 12,13 and 14 increasing the interference and packet drops as well as delay.

Next, we study the impact of the size of the “hole”; i.e. N on the expected per packet delay. Under DIVBAR the packets that arrive at node 2 from source 1 are likely to be forwarded and wander between nodes $12, 13 \dots, 12 + (N - 1)$ before eventually forwarding to 4. In contrast, increasing N has no effect on the performance of D-ORCD. This is because $V_2^7(t) < V_{12+i}^7(t)$, $i = 0, 2, \dots, N - 1$, for all time slots t , in effect, preventing the packets to enter the “hole”. Fig. 7 provides the expected delay encountered by the source packets under various routing policies, as the size of the “hole”, N , increases and the arrival rate is set to low value of $\lambda_1 = 200$ kbps. The figure shows that the average delay under D-ORCD is significantly lower than other candidate protocols as N increases from 1 to 5.

C. Performance of D-ORCD: Grid Topology

We perform simulations for the grid networks of 16 nodes in Fig. 8(a) and 8(b). UDP Traffic is injected at each node $i \in \Omega$, with poison distributed packet arrivals. Figure 9(a) shows the expected delay versus the arrival rate under various routing policies for the network in Fig. 8(a). Under ExOR, packets are always routed opportunistically along the “shortest path” to the destination which results in high delay under heavy traffic scenarios. On the other hand, DIVBAR, E-DIVBAR, and D-ORCD are throughput optimal and hence, they distribute the traffic to ensure bounded average delay for all traffic

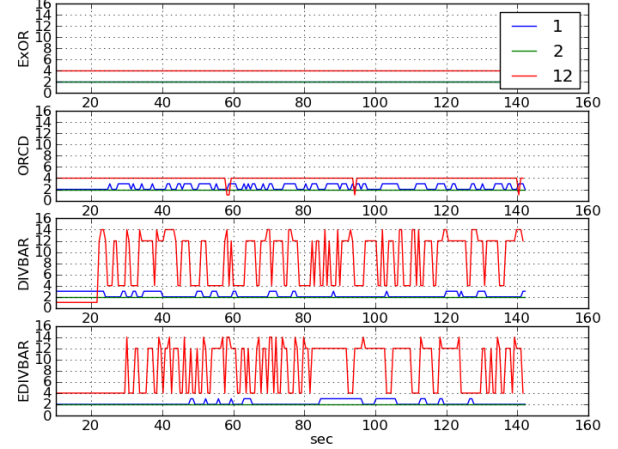
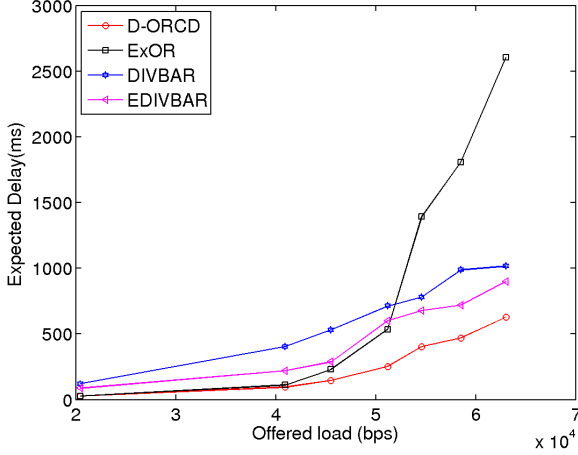


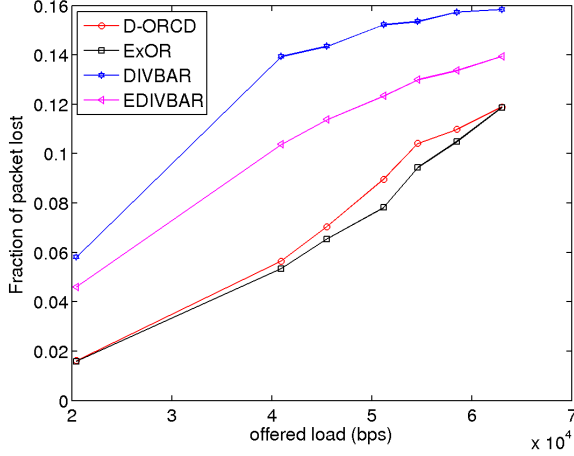
Fig. 6. Highest priority nodes for Canonical Example.

rates inside the stability region. The performance gap between DIVBAR and E-DIVBAR follows from the fact that DIVBAR does not use any metric of closeness to the destination when routing the packets; while E-DIVBAR takes into account the ETX of the nodes. A more interesting observation is the comparable performance of D-ORCD and E-DIVBAR. In other words, in network 8(a) the mere addition of ETX and queue measures in E-DIVBAR perform sufficiently well.

Next, we consider the network shown in Fig.8(b), a modification of the network shown in Fig.8(a) in which link qualities are changed due to the existing barriers in the network. Figure 10(a) shows the delay performance of the candidate routing policies for this network as the traffic load varies. Again, ExOR and DIVBAR show large delay. But, unlike in the case of network shown in Fig.8(a), the performance gap between D-ORCD and E-DIVBAR is now rather significant. The reason is that D-ORCD always route packets along the least congested paths to the destination (without assuming the network topology and the arrival traffic). In other words,



(a) Average delay per packet delivery for Network shown in Fig.8(b).



(b) Fraction of the packets lost is dominated by FO packet loss. (Packet loss due to buffer overflow is negligible)

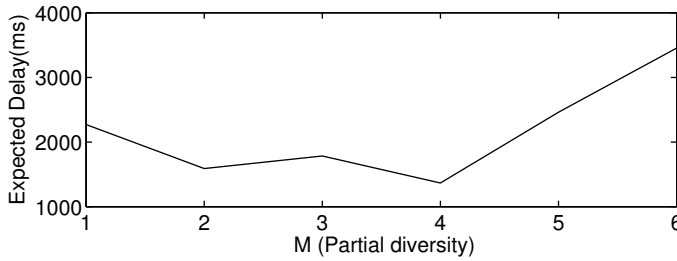


Fig. 10. Delay performance for D-ORCD with partial diversity for Network shown in Fig.8(a)

that computation cycle interval to be sufficiently large. However, to ensure a better delay performance, T_c must be chosen sufficiently low to make the routing decisions more responsive to the instantaneous congestion. In particular, as T_c increases, the chosen routing paths i) utilizes outdated queue lengths and ii) keeps the routing policy fixed for longer durations independent of current queue-lengths. In Figure 11, we plot the performance of D-ORCD as T_c varies in terms of multiple of T_s . We observe that for high load, the choice $T_c = T_s$ outperforms other values for T_c . We have chosen a more

responsive version of (2) at the cost of provable throughput optimality, where T_c is set to $T_s = 0.5$ seconds.

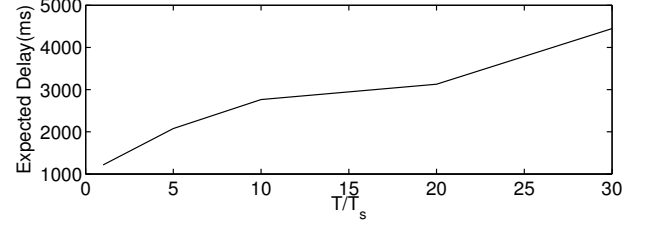


Fig. 11. Delay Performance for D-ORCD for Network shown in Fig.8(a) as T_c varies

V. THEORETICAL GUARANTEES

In this section, we provide a theoretical guarantee regarding the throughput optimality of D-ORCD under the assumptions that i) the flows in the network are destined for the single destination node D (for multi-destination extensions see [23]), ii) link probabilities are time invariant iii) the routing decisions and the successful reception at set S due to transmission from node i is acknowledged perfectly to node i .

Before we precisely state the optimality, we define few notations. We define a *routing decision* $\mu_{ij}(t)$ to be the number of packets (upto 1 packet) whose relaying responsibility is shifted from node i to node j during time slot t ($\mu_{ii}(t) = 1$ means that i retains the packet). Note that $\mu_{ij}(t)$ forms the departure process from node i , while it creates an endogenous arrival to node j . Without loss of optimality, we assume that $p_{ii} = 1$ and $\mu_{iD}(t) = 1$, if $D \in S_i(t)$.

Definition 1. A *routing policy* is a collection of causal routing decisions $\cup_{i,j \in \Omega} \cup_{t=0}^{\infty} \{\mu_{ij}(t)\}$.

Let $A_i(t)$ represent the amount of data that exogenously arrives to node i during time slot t . Arrivals are assumed to be i.i.d. over time and bounded by a constant A_{max} . Let $\lambda_i = \mathbb{E}[A_i(t)]$ denote the exogenous arrival rate to node i . We define $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_N]$ to be the arrival rate vector. Let $Q_i(t)$ denote the queue backlog of node i at time slot t . We assume any data that is successfully delivered to the destination D will exit the network and hence, $Q_D(t) = 0$ for all time slots t . We define $\mathbf{Q}(t) = [Q_1(t), Q_2(t), \dots, Q_{D-1}(t)]$ to be the vector of queue backlogs of nodes $1, 2, \dots, D-1$.

The selection of routing decisions under a routing policy Π together with the exogenous arrivals impact the queue backlog of node i , $i \in \Omega$ as:

$$\begin{aligned} Q_i^\Pi(t+1) &= [Q_i^\Pi(t) - \sum_{j \in \Omega} \mu_{ij}^\Pi(t)]^+ \\ &\quad + \sum_{j \in \Omega} \mu_{ji}^\Pi(t) \mathbf{1}_{\{Q_j^\Pi(t) \geq \mu_{ji}^\Pi(t)\}} + A_i(t), \end{aligned}$$

where the superscript Π emphasizes the dependence of queue backlog dynamics on the choice of policy Π .

Definition 2. Given an ergodic exogenous arrival process with rate λ , a routing policy Π is said to *stabilize* the network if $Q_{tot}^\Pi(t)$ is ergodic and $\mathbb{E}[Q_{tot}^\Pi(t)]$ remains bounded when

packets are routed according to Π . The *stability region* of the network (denoted by \mathfrak{S}) is the set of all arrival rate vectors λ for which there exists a routing policy that stabilizes the network.

Definition 3. A routing policy is said to be *throughput optimal* if it stabilizes the network for all arrival rate vectors that belong to the interior of the stability region.

Fact 1 (Corollary 1 in [7]). An arrival rate vector λ is within the stability region \mathfrak{S} if and only if there exists a stationary randomized routing policy that makes routing decisions $\{\tilde{\mu}_{ij}(t)\}_{i,j \in \Omega}$, solely based on the collection of potential forwarders at time t , $\{S_i(t)\}_{i \in \Omega}$, and for which

$$\mathbb{E} \left[\sum_{j \in \Omega} \tilde{\mu}_{kj}(t) - \sum_{i \in \Omega} \tilde{\mu}_{ik}(t) \right] \geq \lambda_k.$$

We are ready to present Theorem 1 regarding the optimality of D-ORCD.

Theorem 1. Suppose $T_c = O(D)$ and $M = D$. Then D-ORCD is throughput optimal.

The proof of Theorem 1 is based on the Foster-Lyapunov Theorem. For completeness, the structure of the Lyapunov function and a sketch of the proof is provided in the Appendix. By Theorem 1, under D-ORCD, the average total queue backlog remains bounded. Little's theorem implies that under D-ORCD, expected delay is bounded.

Remark 4. Assumptions for optimality of D-ORCD could be relaxed in many cases.

- 1) The packet transmission on a link (i, j) is assumed to be successful with probability p_{ij} , and transmissions on links were assumed to be independent of each other. The computations in (2) and (14) can be generalized to incorporate correlated link qualities. by replacing the term $(\prod_{k \in S} p_{ik}) (\prod_{l \notin S} (1 - p_{il}))$ with $P(S|i)$ in the definition of $P^{(i,d)}(t)$ and $P_{succ-k}^{(i,d)}(t)$, where $P(S|i)$ denotes the probability of the event $\{S_i(t) = S\}$. Furthermore, it is straight forward to show that the throughput optimality of D-ORCD is robust to all channel estimation errors, even though, erroneous link models, in general, can significantly degrade its delay performance.
- 2) In this paper, we assumed that the network topology and the probability of successful transmissions are time-invariant. The generalization to the case of time-varying network topology with stationary transmission probabilities is straight forward [24].

VI. CONCLUSIONS AND DISCUSSIONS

In this paper, combining the important aspects of shortest path routing with those of backpressure routing, we provided a distributed opportunistic routing policy with congestion diversity (D-ORCD) is proposed under which packets are routed according to a rank ordering of the nodes based on a congestion cost measure. Furthermore, we show that D-ORCD allows for a practical distributed and asynchronous 802.11 compatible implementation, whose performance was

investigated via a detailed set of QualNet simulations under practical and realistic networks. Simulations show that D-ORCD consistently outperforms existing routing algorithms in practical settings.

In D-ORCD, we do not model the interference from the nodes in the network, but instead leave that issue to a classical MAC operation. However, the generalization to the networks with inter-channel interference follows directly from [7]. The price of this generalization is shown to be the centralization of the routing/scheduling globally across the network or a constant factor performance loss of the distributed variants [7], [10], [25]. In future, we are interested in generalising D-ORCD for joint routing and scheduling optimizations as well consider system level implications. Incorporating throughput optimal CSMA based MAC scheduler (proposed in [26]) with congestion aware routing is also promising area of research.

The design of D-ORCD requires knowledge of channel statistics. Designing congestion control routing algorithms to minimize expected delay without the topology and the channel statistics knowledge is an area of future research.

ACKNOWLEDGEMENTS

This work was partially supported by the industrial sponsors of UCSD Center for Wireless Communications (CWC) and Center for Networked Systems (CNS) and UC Discovery Grant #com07-10241. The authors would like to thank Mr. Anders Plymoth for valuable discussions and suggestions.

APPENDIX

We provide a sketch of the proof for the throughput optimality of D-ORCD for a connected network.³

A. Relationship to Centralized ORCD

We prove the throughput optimality by relating D-ORCD update equation (2) to the convergence of closely related fixed point equation. In particular, we relate the routing decisions for D-ORCD with the decisions taken according to the congestion measures $\{V_i^*(t)\}$ obtained from the fixed point equation:

$$V_i^*(t) = \frac{Q_i(t)}{P^{(i,D)}(t)} + \sum_{k: H^{(i,D)}(t)} \frac{P_{succ-k}^{(i,D)}(t)}{P^{(i,D)}(t)} V_k^*(t). \quad (14)$$

We refer to the centralized routing algorithm which makes decisions at each instant according to $V_i^*(t)$ as C-ORCD. Following lemma states a relationship between D-ORCD and C-ORCD.

Lemma 1. Assume T_c is sufficiently large ($T_c \sim O(D)$). Then during $T(t) \leq t < T(t + T_c)$, (2) converges to the fixed point equation (14), i.e. $\{V_i^D(T(t + T_c))\}_{i \in \Omega}$ solves (14) and $V_i^D(T(t + T_c)) = V_i^*(T(t))$.

Proof: The convergence $V_i^D(t) \rightarrow V_i^*(T(t))$ during $T(t) \leq t < T(t + T_c)$ follows by relating (2) to the Bellman-Ford algorithm with fixed link cost. It is known from

³In connected network each node has a positive probability path to the destination. If a node has no path to the destination, it cannot sustain any traffic and can be ignored without loss of generality.

[27, Theorem 2.4] that asynchronous distributed Bellman-Ford algorithm converges in finite time when the control packets are instantaneously received (or control packets are timestamped and older packets are discarded). Furthermore, with high probability, the time until the termination of the asynchronous Bellman-Ford algorithm is $O(D)$, where D is the number of nodes in the network. ■

Note that when implementing D-ORCD, we broadcast control packets using high priority and the control packets do not undergo backoff. This ensures that with high probability the packets are instantaneously received. Thus the convergence in Lemma 1 is justified.

We now provide the proof for the throughput optimality of D-ORCD for $T_c \sim O(D)$. To simplify the notations, let policies in C-ORCD and D-ORCD be denoted by π^* and $\hat{\pi}$ respectively. In [13], the authors constructed an appropriate Lyapunov function L^* to show that C-ORCD is throughput optimal. In particular it is shown that:

Fact 2. There exists a Lyapunov function. $L^* : \mathbb{R}_+^D \rightarrow \mathbb{R}_+$ such that for all time slots t and $B > 0$, $\epsilon > 0$,

$$\mathbb{E} [L^*(Q^{\pi^*}(t+1)) - L^*(Q(t)) | Q(t)] \leq B - \epsilon \sum_{k=1}^N Q_k(t), \quad (15)$$

where superscript π^* in Q^{π^*} implies the dependence of backlog vector on the routing policy π^* .

To prove throughput optimality of D-ORCD, it suffices to show that the Lyapunov drifts under π^* and $\hat{\pi}$ have a bounded difference. More precisely,

Lemma 2. Let L^* be the Lyapunov function as proposed in [13]. Then for $B'' > 0$,

$$\mathbb{E} [L^*(Q^{\hat{\pi}}(t+1)) | Q(t)] - \mathbb{E} [L^*(Q^{\pi^*}(t+1)) | Q(t)] < B''.$$

Lemma 2 together with (15) implies the existence of Lyapunov function with negative expected drift i.e. for Lyapunov function L^* , there exists $B' > 0$, $\epsilon' > 0$ such that,

$$\mathbb{E} [L^*(Q^{\hat{\pi}}(t+1)) - L^*(Q^{\pi^*}(t)) | Q(t)] \leq B' - \epsilon' \sum_{k=1}^N Q_k(t).$$

The details of the construction of L^* and the proof of Lemma 2 is provided in Appendix C.

B. Review of C-ORCD results

Before proceeding, we introduce some notations. Let $[x]^+ = \max\{x, 0\}$. The indicator function $\mathbf{1}_{\{X\}}$ takes the value 1 whenever event X occurs, and 0 otherwise. For any set S , $|S|$ denotes the cardinality of S , while for any vector \mathbf{v} , $\|\mathbf{v}\|$ denotes the euclidean norm of \mathbf{v} . When dealing with a sequence of sets C_1, C_2, \dots , we define $C^i = \cup_{j=1}^i C_j$.

The following definitions are required in order to identify the Lyapunov functions for C-ORCD and D-ORCD.

Definition 4. A rank ordering $R = (C_1, C_2, \dots, C_M)$ is an ordered list of non-empty sets C_1, C_2, \dots, C_M ($1 \leq M \leq D$), referred to as *ranking classes*, that make up a partition of the

set of nodes $\{1, 2, \dots, D\}$, i.e., $\cup_{i=1}^M C_i = \{1, 2, \dots, D\}$ and $C_i \cap C_j = \emptyset$, $i \neq j$.

Definition 5. A rank ordering $R = (C_1, C_2, \dots, C_M)$ is referred to as *path-connected* if for each node $i \in C_k$, $1 \leq k \leq M$, there exist distinct nodes $j_1, j_2, \dots, j_l \in C^{k-1}$ such that $p_{ij_1} > 0, p_{j_1j_2} > 0, \dots, p_{j_lD} > 0$. The set of all path-connected rank orderings is denoted by \mathcal{R}^c .

Let f be a bivariate function of the following form:

$$f(m, n) = \frac{1}{K^m(K^n - 1)} \quad \text{for all } m \geq 0, n > 0,$$

where $K = 1 + \frac{1}{p_{\min}}$ for $p_{\min} = \min\{p_{ij} : i, j \in \Omega, p_{ij} > 0\}$.

In [13], the authors proposed a method that utilizes the bivariate function f and partitions the space of queue backlogs, \mathbb{R}_+^D , into $|\mathcal{R}^c|$ cones denoted by $\{D_f^c(R)\}_{R \in \mathcal{R}^c}$. The piecewise Lyapunov function, $L_f^* : \mathbb{R}_+^D \rightarrow \mathbb{R}_+$, is then constructed by assigning to each cone $D_f^c(R)$, $R = (C_1, C_2, \dots, C_M) \in \mathcal{R}^c$, a weighted quadratic function of the form:

$$L_f(Q, R) = \sum_{i=1}^M f(|C^{i-1}|, |C_i|) Q_{C_i}^2, \quad (16)$$

where $Q_C(t) = \sum_{i \in C} Q_i(t)$. More precisely,

$$L_f^*(Q) = \sum_{R \in \mathcal{R}^c} L_f(Q, R) \mathbf{1}_{\{Q \in D_f^c(R)\}}. \quad (17)$$

Let us consider the Lyapunov drift when $Q(t) \in D_f^c(R)$ for some $R = (C_1, C_2, \dots, C_M) \in \mathcal{R}^c$. Since the collection of cones $\{D_f^c(R)\}_{R \in \mathcal{R}^c}$ partitions \mathbb{R}_+^D , we define function $U_f : \Omega \times \mathbb{R}_+^D \rightarrow \mathbb{R}_+$ such that

$$U_f(k, Q) = f(|C^{i-1}|, |C_i|) Q_{C_i}(t), \quad (18)$$

where $Q \in D_f^c(R)$, $R = (C_1, C_2, \dots, C_M)$, and $k \in C_i$.

Let $\{\mu_{ij}^*\}_{i,j \in \Omega}$ represent the routing decisions under C-ORCD, while $\{\hat{\mu}_{ij}(t)\}_{i,j \in \Omega}$ represent the routing decisions under D-ORCD. For ease of notation and exposition define $A_C(t) = \sum_{i \in C} A_i(t)$,

$$\mu_{C,in}(t) = \sum_{k=1}^D \sum_{j \in C} \mu_{kj}(t) \mathbf{1}_{\{Q_k(t) \geq 1\}}, \quad \text{and} \\ \mu_{C,out}(t) = \sum_{k \in C} \sum_{j=1}^D \mu_{kj}(t) \mathbf{1}_{\{Q_k(t) \geq 1\}}. \quad \text{We have,}$$

$$\begin{aligned} & L_f^*(Q(t+1)) - L_f^*(Q(t)) \\ & \stackrel{(a)}{=} \sum_{i=1}^M f(|C^{i-1}|, |C_i|) [Q_{C_i}^2(t+1) - Q_{C_i}^2(t)] \\ & \quad + O(\|Q(t+1) - Q(t)\|^2) \\ & \stackrel{(b)}{=} -2 \sum_{i=1}^M f(|C^{i-1}|, |C_i|) Q_{C_i}(t) (\mu_{C_i,out}(t) - \mu_{C_i,in}(t) \\ & \quad - A_{C_i}(t)) + O(1) \\ & = -2 \sum_{i=1}^M f(|C^{i-1}|, |C_i|) Q_{C_i}(t) \mathbf{1}_{\{Q_k(t) \geq 1\}} \\ & \quad \left(\sum_{j=1}^D \sum_{k \in C_i} \mu_{kj}(t) - \sum_{k=1}^D \sum_{j \in C_i} \mu_{kj}(t) - A_{C_i}(t) \right) + O(1), \end{aligned} \quad (19)$$

where (a) follows from continuity and differentiability of L_f^* [13, Lemma 3] and writing $L_f^*(\mathbf{Q}(t+1))$ in terms of its first-order Taylor expansion around $L_f^*(\mathbf{Q}(t))$, and (b) follows from Fact 3 below.

Fact 3 ([13]). Let $R = (C_1, C_2, \dots, C_M) \in \mathcal{R}$ and $\mathbf{Q}(t) \in D_f(R)$. We have

$$Q_{C_i}^2(t+1) - Q_{C_i}^2(t) = \beta_f - 2Q_{C_i}(t)(\mu_{C_i, \text{out}}(t) - \mu_{C_i, \text{in}}(t) - A_{C_i}(t)),$$

where β_f is a constant bounded real number.

Finally from (18) and (19),

$$\begin{aligned} & L_f^*(\mathbf{Q}(t+1)) - L_f^*(\mathbf{Q}(t)) \\ &= -2 \sum_{i=1}^M \sum_{k \in C_i} \sum_{l=1}^D U_f(k, Q) \mu_{kl}(t) \mathbf{1}_{\{Q_k(t) \geq 1\}} \\ & \quad - \sum_{i=1}^M \sum_{l \in C_i} \sum_{k=1}^D U_f(l, Q) \mu_{kl}(t) \mathbf{1}_{\{Q_k(t) \geq 1\}} - \sum_{i=1}^M A_{C_i}(t) + O(1) \\ &= -2 \sum_{k=1}^D \sum_{l=1}^D U_f(k, Q) \mu_{kl}(t) \mathbf{1}_{\{Q_k(t) \geq 1\}} \\ & \quad - \sum_{k=1}^D \sum_{l=1}^D U_f(l, Q) \mu_{kl}(t) \mathbf{1}_{\{Q_k(t) \geq 1\}} - \sum_{i=1}^M A_{C_i}(t) + O(1). \end{aligned} \quad (20)$$

Fact 4 ([13]). Routing decisions under C-ORCD are such that $\mu_{ij}^* = 1$, only when $j \in S_i(t)$ and $U_f(j, \mathbf{Q}(t)) \leq U_f(k, \mathbf{Q}(t))$ for all $k \in S_i(t)$.

This fact together with (20) provides the proof of Fact 2.

C. Proof of Lemma 2

We begin the proof of Lemma 2 by stating following Claim.

Claim 1. Routing decisions under D-ORCD are such that $\hat{\mu}_{ij} = 1$, only when $j \in S_i(t)$ and $U_f(j, \hat{\mathbf{Q}}(t)) \leq U_f(k, \hat{\mathbf{Q}}(t))$ for all $k \in S_i(t)$, where $\hat{\mathbf{Q}}(t) = \bar{\mathbf{Q}}(T(t))$.

With this, we are ready to proceed with the proof of Lemma 2.

$$\begin{aligned} & \mathbb{E}[L_f^*(\mathbf{Q}^{\hat{\pi}}(t+1)) | \mathbf{Q}(t)] - \mathbb{E}[L_f^*(\mathbf{Q}^{\pi^*}(t+1)) | \mathbf{Q}(t)] \\ &= \mathbb{E} \left[2 \sum_{k=1}^D \sum_{l=1}^D (\hat{\mu}_{kl}(t) - \mu_{kl}^*(t)) \left(U_f(k, \mathbf{Q}(t)) \right. \right. \\ & \quad \left. \left. - U_f(l, \mathbf{Q}(t)) \mathbf{1}_{\{Q_k(t) \geq 1\}} \right) | \mathbf{Q}(t) \right] + O(1), \end{aligned} \quad (21)$$

where equality follows from (20).

Suppose node i 's transmission at time t is received by potential forwarders $S_i(t)$. Furthermore, suppose that nodes $a, b \in S_i(t)$ are the nodes with the highest rank under C-ORCD and D-ORCD respectively, i.e. $\mu_{ia}^*(t) = \hat{\mu}(t)_{ib} = 1$. From Fact 4 and Claim 1, we have

$$U_f(a, \mathbf{Q}(t)) \leq U_f(b, \mathbf{Q}(t)), \quad (22)$$

$$U_f(a, \hat{\mathbf{Q}}(t)) \geq U_f(b, \hat{\mathbf{Q}}(t)). \quad (23)$$

In order to prove Lemma 2 it suffices to show that

$$U_f(b, \mathbf{Q}(t)) - U_f(a, \mathbf{Q}(t)) = O(\|\mathbf{Q}(t) - \hat{\mathbf{Q}}(t)\|). \quad (24)$$

Consider the line that connects $\mathbf{Q}(t)$ and $\hat{\mathbf{Q}}(t)$ in \mathbb{R}_+^D . Suppose this line goes through $M-1$ cones in \mathbb{R}_+^D . Let Z_1, Z_2, \dots, Z_M be respectively the intersection of the line connecting $\mathbf{Q}(t)$ to $\hat{\mathbf{Q}}(t)$ with the M separating hyperplanes of the $M-1$ cones between them, i.e.

$$\begin{aligned} & \|\mathbf{Q}(t) - \hat{\mathbf{Q}}(t)\| \\ &= \|\mathbf{Q}(t) - \mathbf{Z}_1\| + \|\mathbf{Z}_1 - \mathbf{Z}_2\| + \dots + \|\mathbf{Z}_M - \hat{\mathbf{Q}}(t)\|. \end{aligned}$$

Note that since Z_1, Z_2, \dots, Z_M are on the hyperplanes, every two consecutive points in set $\{\mathbf{Q}(t), Z_1, Z_2, \dots, Z_M, \hat{\mathbf{Q}}(t)\}$ can be considered to belong to the same cone, and hence, have same rank ordering of the nodes. From definition of function U_f , we obtain

$$\begin{aligned} & |U_f(a, \mathbf{Q}(t)) - U_f(a, \mathbf{Z}_1)| = O(\|\mathbf{Q}(t) - \mathbf{Z}_1\|), \\ & |U_f(a, \mathbf{Z}_m) - U_f(a, \mathbf{Z}_{m+1})| = O(\|\mathbf{Z}_m - \mathbf{Z}_{m+1}\|), \quad 1 \leq m \leq M, \\ & |U_f(a, \mathbf{Z}_M) - U_f(a, \hat{\mathbf{Q}}(t))| = O(\|\mathbf{Z}_M - \hat{\mathbf{Q}}(t)\|). \end{aligned}$$

Therefore,

$$\begin{aligned} & U_f(a, \mathbf{Q}(t)) - U_f(a, \hat{\mathbf{Q}}(t)) \\ &= [U_f(a, \mathbf{Q}(t)) - U_f(a, \mathbf{Z}_1)] + [U_f(a, \mathbf{Z}_1) - U_f(a, \mathbf{Z}_2)] \\ & \quad + \dots + [U_f(a, \mathbf{Z}_M) - U_f(a, \hat{\mathbf{Q}}(t))] = O(\|\mathbf{Q}(t) - \hat{\mathbf{Q}}(t)\|). \end{aligned}$$

We can derive the same result for all other nodes in the network. In other words, there exist constants η_a, η_b such that

$$U_f(a, \mathbf{Q}(t)) = U_f(a, \hat{\mathbf{Q}}(t)) + \eta_a \|\mathbf{Q}(t) - \hat{\mathbf{Q}}(t)\|, \quad (25)$$

$$U_f(b, \mathbf{Q}(t)) = U_f(b, \hat{\mathbf{Q}}(t)) + \eta_b \|\mathbf{Q}(t) - \hat{\mathbf{Q}}(t)\|. \quad (26)$$

However, (22), (23), (25), and (26) imply that $\eta_a \leq \eta_b$ and

$$U_f(b, \mathbf{Q}(t)) - U_f(a, \mathbf{Q}(t)) \leq (\eta_b - \eta_a) \|\mathbf{Q}(t) - \hat{\mathbf{Q}}(t)\|. \quad (27)$$

With this, the proof is now complete.

REFERENCES

- [1] P. Larsson, "Selection Diversity Forwarding in a Multihop Packet Radio Network with Fading channel and Capture," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 2, no. 4, pp. 47–54, October 2001.
- [2] M. Zorzi and R. R. Rao, "Geographic Random Forwarding (GeRaF) for Ad Hoc and Sensor Networks: Multihop Performance," *IEEE Transactions on Mobile Computing*, vol. 2, no. 4, 2003.
- [3] S. Biswas and R. Morris, "ExOR: Opportunistic Multi-hop Routing for Wireless Networks," *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 33–44, October 2005.
- [4] C. Lott and D. Teneketzis, "Stochastic routing in ad-hoc networks," *IEEE Transactions on Automatic Control*, vol. 51, pp. 52–72, January 2006.
- [5] S.R. Das S. Jain, "Exploiting path diversity in the link layer in wireless ad hoc networks," *World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a*, pp. 22–30, June 2005.
- [6] Parul Gupta and Tara Javidi, "Towards Throughput and Delay Optimal Routing for Wireless Ad-Hoc Networks," in *Asilomar Conference*, November 2007, pp. 249–254.
- [7] M. J. Neely and R. Ugaonkar, "Optimal backpressure Routing for Wireless Networks with Multi-Receiver Diversity," *IEEE Transactions on Information Theory*, pp. 862–881, July 2009.

- [8] L. Tassiulas and A. Ephremides, "Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1949, August 1992.
- [9] S. Sarkar and S. Ray, "Arbitrary Throughput Versus Complexity Trade-offs in Wireless Networks using Graph Partitioning," *IEEE Transactions on Automatic Control*, vol. 53, no. 10, pp. 2307–2323, November 2008.
- [10] Y. Xi and E. M. Yeh, "Throughput Optimal Distributed Control of Stochastic Wireless Networks," in *WiOpt*, April 2006.
- [11] B. Smith and B. Hassibi, "Wireless Erasure Networks with Feedback," arXiv:0804.4298v1, 2008.
- [12] Y. Yi and S. Shakkottai, "Hop-by-Hop Congestion Control Over a Wireless Multi-Hop Network," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 133–144, February 2007.
- [13] M. Naghshvar, H. Zhuang, and T. Javidi, "A General Class of Throughput Optimal Routing Policies in Multi-hop Wireless Networks," 2009, arXiv:0908.1273v1.
- [14] E. Leonardi, M. Mellia, M. A. Marsan, and F. Neri, "Optimal Scheduling and Routing for Maximum Network Throughput," *IEEE/ACM Transactions on Networking*, vol. 15, no. 6, pp. 1541–1554, December 2007.
- [15] S. Shakkottai, L. Ying, and A. Reddy, "On Combining Shortest-Path and Back-Pressure Routing Over Multihop Wireless Networks," *IEEE/ACM Transactions on Networking*, vol. 19, June 2011.
- [16] L. Huang, S. Moeller, M. J. Neely, and B. Krishnamachari, "LIFO-Backpressure Achieves Near Optimal Utility-Delay Tradeoff," in *Proc. of 9th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 2011.
- [17] W. Stallings, *Wireless Communications and Networks*, Prentice Hall, second edition, 2004.
- [18] Aman Shaikh, Anujan Varma, Lampros Kalampoukas, and Rohit Dube, "Routing stability in congested networks: Experimentation and analysis," in *In Proc. ACM SIGCOMM*, 2000, pp. 163–174.
- [19] J. M. Jaffe and F. Moss, "A Responsive Distributed Routing Algorithm for Computer Networks," *IEEE Transactions on Communications*, vol. 30, July 1982.
- [20] Uysell Black, *IP routing protocols: RIP, OSPF, BGP, PNNI and Cisco routing protocols*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 2000.
- [21] J. Doble, *Introduction to Radio Propagation for Fixed and Mobile Communications*, Artech House, Boston, 1996.
- [22] M. J. Neely, "Optimal Backpressure Routing for Wireless Networks with Multi-Receiver Diversity," in *Conference on Information Sciences and Systems (CISS)*, March 2006.
- [23] Hairuo Zhuang, *Throughput optimal routing in wireless Ad-hoc networks*, Ph.D. thesis, University of California, San Diego, 2010.
- [24] M. Naghshvar and T. Javidi, "Opportunistic Routing with Congestion Diversity in Wireless Multi-hop Networks," *IEEE INFOCOM*, March 2010.
- [25] L. Ying and S. Shakkottai, "on throughput-optimal scheduling with delayed channel state feedback," in *Information Theory and Applications Workshop*, 2008.
- [26] Libin Jiang and J. Walrand, "A Distributed CSMA Algorithm for Throughput and Utility Maximization in Wireless Networks," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, Sept. 2008, pp. 1511–1519.
- [27] J. N. Tsitsiklis and G. D. Stamoulis, "On the Average Communication Complexity of Asynchronous Distributed Algorithms," *Journal of the ACM*, vol. 42, pp. 382 – 400, 1995.